# SPECTRAL IMAGE PROCESSING AND ANALYSIS OF THE ARCHIMEDES PALIMPSEST

Roger L. Easton, Jr., William A. Christens-Barry, Keith T. Knox

Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester NY, USA
Rochester NY, 14623-5604, USA 1-585-475-5969, 1-585-475-5988 (fax), e-mail: *easton@cis.rit.edu*
web: *www.cis.rit.edu/people/faculty/easton*

## ABSTRACT

*The Archimedes Palimpsest is a $10^{th}$-century parchment manuscript that was erased in the $13^{th}$ century and overwritten with a Christian prayer book. The Walters Art Museum in Baltimore, MD, USA has supervised a ten-year collaboration by conservators, imaging scientists, and scholars to image and transcribe the original writings. This paper reports on the variety of optical imaging and image processing techniques used in the project. Image analysis tools to clarify texts included supervised segmentation via a spectral pseudoinverse calculation, deterministic renderings in pseudocolor, and statistical analysis with dynamical pseudocolor rendering.*

## 1. SIGNIFICANCE OF THE CODEX

Almost everything known about the work of Archimedes has been gleaned from three codex manuscripts. The first two vanished from scholarly view by the $16^{th}$ century, but parts of seven treatises survive in the third, a Byzantine codex copied in the $10^{th}$ century. It includes the only extant leaves of *On the Method of Mechanical Theorems*, where Archimedes used mechanical analogies to prove mathematical theorems, the only known leaf of *Stomachion*, and the only copy of *On Floating Bodies* in the original Greek. The treatises were originally copied in a large format, with page size approximately 200mm × 300mm. This codex and pages from other important books were sacrificed in 1229 CE to make a copy of an Orthodox Christian liturgical book, the *Euchologion*; the pages were disbound, the text erased, each folio cut along the fold, and the prayer book copied over the newly cleaned pages perpendicular to the original writing. Recycling of books in this manner was a common practice due to the expense of making new parchment. Where visible, the iron gall ink of the erased texts appears "reddish," while the characters of the later prayer book are dark brown in color and most remain quite readable.

A discussion of the mysterious path followed by the manuscript from 1229 CE to the present day is beyond the scope of this paper and is covered in detail elsewhere [1]. Suffice it to say that the codex was studied in the early 1900s by Johan Ludvig Heiberg and not seen again publicly until auctioned by Christie's in 1998. During that span of nearly a century, many pages were damaged severely by mold, some disappeared altogether, and four pages were covered by painted portraits of Christian Evangelists. The manuscript was purchased at auction for $2 million US by an anonymous American collector, who deposited it with the Walters Art Museum in Baltimore, MD, USA and funded an intensive program of conservation, imaging, and scholarly study. The imaging phase was officially completed on October 28, 2008, the tenth anniversary of the auction, when all original and many processed images were posted on the project website for free use by other researchers under a Creative Commons attribution license [2].

This paper reports on the range of spectral reflectance and fluorescence optical imaging techniques used in the transcription of the texts. X-ray fluorescence (XRF) imaging, which proved essential on the overpainted leaves and the colophon of the *Euchologion*, is considered elsewhere [3]. Among the exciting aspects of this project were the discoveries of additional unique works in the manuscript: parts of two speeches by the Athenian orator Hypereides and of a commentary on Aristotle's "Categories" by Alexander of Aphrodisias, a partial history of St. Pantaleon, and the *Menaion*, an orthodox liturgical text for the $10^{th}$ century. Because of differences in the conditions of the pages and the inks, different imaging and processing techniques were used to recover the original texts from these pages.

## 2. SPECTRAL AND FLUORESCENCE IMAGING

The visual differences in color of the overtext, undertext, and parchment suggested the use of spectral imaging to recover the original writings. Though the technique is often further distinguished as "multispectral" or "hyperspectral," the arbitrariness of these classes encourages the use of the more general term. In spectral imaging, digital images of the same scene are collected in different wavebands and combined by computer in various ways to enhance the desired feature(s) of the scene. A key concept of spectral image processing is that images at different wavelengths are *combined* arithmetically to recover the text, unlike the simpler tactic of selecting the single spectral band from those available where the feature of interest is most visible, which is sometimes described as multispectral image processing [4]. The basic techniques of

# Report Documentation Page

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE | 2. REPORT TYPE | 3. DATES COVERED |
|---|---|---|
| **2011** | **N/A** | **-** |

4. TITLE AND SUBTITLE
**Spectral Image Processing And Analysis Of The Archimedes Palimpsest**

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
**Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester NY, USA**

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
**Approved for public release, distribution unlimited**

13. SUPPLEMENTARY NOTES
**See also ADA589869. European Signal Processing Conference (19th) (EUSIPCO) Held in Barcelona, Spain on August 29-September 2, 2011.**

14. ABSTRACT
**The Archimedes Palimpsest is a 10th-century parchment manuscript that was erased in the 13th century and overwritten with a Christian prayer book. The Walters Art Museum in Baltimore, MD, USA has supervised a ten-year collaboration by conservators, imaging scientists, and scholars to image and transcribe the original writings. This paper reports on the variety of optical imaging and image processing techniques used in the project. Image analysis tools to clarify texts included supervised segmentation via a spectral pseudoinverse calculation, deterministic renderings in pseudocolor, and statistical analysis with dynamical pseudocolor rendering.**

15. SUBJECT TERMS

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | |
| **unclassified** | **unclassified** | **unclassified** | **SAR** | **5** | |

spectral imaging have a long history in military and environmental remote sensing and have been applied to historical artifacts since scientific digital cameras became widely available in approximately the mid-1990s.

The collected spectral imagery usually provides a measurement of spectral reflectance, but it also is very useful to obtain imagery of fluorescence emitted after absorption of light photons with higher energies (often in the ultraviolet spectrum); the spectrum of the emitted fluorescence is characteristic of the material. Under incident ultraviolet light, the fluorescence of parchment is dominated by blue visible light that can be imaged with a standard silicon sensor, though we shall see that very useful information about the original texts may be conveyed by the fluorescence spectrum

The original exploratory imaging sessions took place in the summer of 2000 and led to adoption of a standard imaging protocol. The team assembled at the Walters Art Museum six times between 2001 and 2004 to image approximately 15 newly disbound and conserved folios. An additional session was held in November 2006 to test several experimental techniques. The entire codex was reimaged in August 2007 with a higher-resolution camera under spectral illumination from light-emitting diodes, and some pages were imaged yet again in March 2009 to provide data for new processing algorithms that had been developed.

Over the course of the project, one important lesson learned (and reinforced in work on other manuscripts) is that every method implemented has been useful, if not essential, for some page. In short, it is essential to have a variety of hardware and software tools in the arsenal and to be flexible in their application.

## 3. PHASE-I IMAGING (2000)

The original goal of the imaging in the summer of 2000 was to develop a scheme to collect and process spectral imagery to "strip off" the later writings and leave the original "undertext" with enhanced contrast and readability; this objective is perhaps better described as making the overtext "disappear" into the parchment to improve the visibility of the original text. After considerable experimentation, a method was developed that works well for the Archimedes texts and for leaves from some of the other manuscripts, though the protocol for imaging and processing had to be modified for use on those other leaves. The method required significant custom processing.

The original method collected images with a 12-bit scientific camera in six broad spectral bands from the ultraviolet through the near-infrared. These were combined using the standard processing algorithm of supervised least-squares spectral classification [5] to recover images of each class of object (parchment, original ink, later ink, etc.). The example in Figure 1 compares the original appearance to processed images of Euchologion leaf f.70v for the two text classes. Though the images required time-consuming custom processing to register the image sets, the imaging team was pleased with these results and believed that they validated the proposed protocol. In addition, it provided a systematic tool that could be applied to spectral imaging of manuscripts, an area where *ad hoc* techniques are quite common [6].

An important lesson learned in this phase was that the images must satisfy the scholars and not the imagers. To the surprise of the imagers, the scholars judged the images to be inadequate for scholarly reading, in part because they were perceived as "fuzzy." The assessment of "fuzziness" required substitution of a higher-resolution sensor and a new protocol to eliminate issues with registering the different spectral images. It also forced changes to the original goal of making the overtext disappear into the background parchment. In a seeming paradox, the scholars felt that this made the Archimedes text more difficult to read; rather, they preferred to see both texts but wanted them distinguished by a different property (e.g., color).

(a) (b) (c)

**Figure 1: Original multispectral processing (f.70v): (a) visual appearance; (b) class of Euchologion text rendered as "white;" (c) Archimedes text in "white." Note effect of mold in upper-right quadrant. (Copyright retained by owner of the Archimedes Palimpsest).**

## 4. PHASE-II PRODUCTION IMAGING (2001-2005)

The imaging plan was modified in 2001 to meet a new goal of enhancing the contrast and appearance of the original undertext while retaining visible overtext. The new process was based on the observation that erased text is generally "redder" than later text, so that the original characters may be nearly invisible under red illumination. The neutral color and high contrast of the overtext makes it very evident under all visible colors of illumination. This led to a simple algorithm for enhancing the original undertext from images taken with a color digital camera of the type used by photojournalists (Kodak DCS-760). Its six-megapixel sensor incorporates a Bayer RGB screen so that each pixel measures the light from one of the three additive primary colors. With a 60mm Micro-Nikkor lens, the camera produced images with a spatial resolution of 25 pixels per mm (~ 600 dpi); the individual images covered areas of approximately 120mm × 80mm on the page. Each full folio was imaged in ten overlapping sections under three illuminations: white-light xenon strobe to produce documentary images of the visible appearance, low-wattage reddish tungsten light, and longwave ultraviolet light ($\lambda \sim 365$ nm). To image an entire folio, the manuscript was moved beneath the camera by a computer-controlled x-y translation table. The images collected under the "reddish" tungsten and the ultraviolet lamps were combined to render the processed image. Since the visible wavelengths closest to the ultraviolet excitation are in the blue, light in this channel dominates the fluorescence images, though we shall see that useful fluorescence information exists at longer wavelengths on some leaves. The red channel of the tungsten image shows very little evidence of the reddish original text, while both texts generally are visible in the blue channel of the ultraviolet image. To balance the local contrast, the images were "normalized" based on the statistics of gray values in a neighborhood of each pixel of selectable size (typically 401 pixels square). The output pixel value is based on the difference of the original gray value from the mean and variance in the local neighborhood; the values were scaled to fill ±3 standard deviations about the mean mid-gray value.

The two channels after contrast enhancement were placed in different channels of a "pseudocolor" image; the tungsten-red channel (with "bright" original text and "dark" Euchologion text) in the red channel, while the ultraviolet-blue image (with both texts "dark") in the blue and green channels. The original text appears "bright" in red and "dark" in green and blue and thus appears with a reddish tint, while the Euchologion text is dark in all three and appears in a neutral gray. This color cue provides information to the reader about the origin of the characters (Figure 2).

The processed pseudocolor image sections were digitally "stitched" into large images of complete folios with approximately $5000 \times 7500$ pixels. The image sets were distributed to scholars for their assessment on portable disk drives and/or as prints. The scholars judged that they could read as much as 80% of the original text from these images, and this success prodded them to urge development of techniques for the remaining 20%. [7,8]

**Figure 2: Pseudocolor image of gutter region of f.093v-092r showing increased contrast of differential color rendering of undertext that is oriented horizontally. (Copyright retained by owner of the Archimedes Palimpsest).**

## 5. PHASE-III, REIMAGING WITH LEDS (2007)

The advances in imaging technology persuaded the owner of the palimpsest to fund a complete reimaging of the book using a spectral light-emitting diode (LED) illuminator from .Equipoise Imaging, LLC and a higher-resolution camera supplied by Stokes Imaging.

LEDs generate light from electronic transitions rather than from thermal interactions and thus generate little heat that might damage manuscripts. The narrow wavebands ($\Delta\lambda \sim 25\text{-}40$ nm) also have advantages over broadband illumination for spectral image processing. The original Archimedes LED illumination system (now dubbed "Eureka Lights") constructed by Dr. Christens-Barry included one ultraviolet band, seven visible bands and four infrared bands, plus raking illumination from two sides at blue and infrared wavelengths for those parchments with etched text as just described. These illuminators are now incorporated into the standard imaging system now used in several venues [9], including the U.S. Library of Congress, and projects to image the Oxyrhynchus papyri and the Dead Sea Scrolls.

The original plan called for a camera with a 33-megapixel monochrome sensor, but delays in delivery forced substitution of a Sinar 54H 22-megapixel color digital back that incorporates piezoelectric micropositioners capable of translating the sensor relative to the optical image by full and half-pixel increments along both axes. The full-pixel translations allow measurement of red, green, and blue values at all pixel locations, which eliminated the need for color interpolation at each pixel. The half-pixel translations allow doubling the pixel count along each axis for stationary

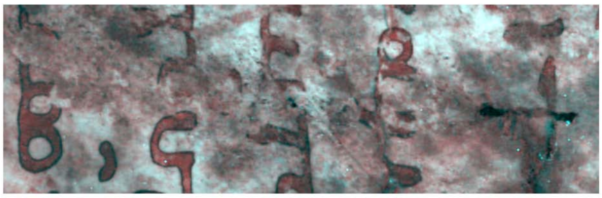
objects to produce 88-megapixel images. Though this capability was utilized, the four-fold increase in collection time (to 32 minutes per leaf) and the much-reduced modulation at the newly acquired large spatial frequencies significantly diminished any value of the additional pixels.

## 6. EXPERIMENTAL IMAGING TECHNIQUES

### 6.1 RAKING-INCIDENCE ILLUMINATION (2006)

One new technique was implemented based on observations by Dr. Judson Hermann of Allegheny University, who was trying to read the fragments of speeches by the Greek orator Hypereides. Dr. Hermann noticed that the original ink of this manuscript had disappeared but that the acid in the ink had left shallow "channels" in the parchment that were more visible if illuminated at raking angles. The imaging system used in the extra session in November 2006 was modified to include a similar illuminator to that used with the microscope. The value is demonstrated in a comparison of the pseudocolor and raking-incidence images in Figure 3.

(a)



(b)

**Figure 3: (a) original pseudocolor rendering of f. 176v showing little apparent undertext; (b) original characters spelling "διονδασ" ("*Diondas*") become visible under raking illumination (Copyright retained by the owner of the Archimedes Palimpsest)**.

### 6.2 PCA OF RGB FLUORESCENCE (2007-2008)

The use of the color sensor was originally a disappointment, but later proved to be valuable as the color information at each pixel led to some new text discoveries on the pages of the commentary on Aristotle. This text was first identified in June 2005, when Nigel Wilson of Lincoln College of Oxford University was able to read the Greek characters for "Aristotle" in the gutter of f.80r. Unfortunately, pseudocolor processing was of very little help on the folios of this text. A method to read these texts from the color images was developed in 2008, thanks to a fortuitous convergence of events. In response to a challenge by Dr. Noel, Kevin Bloechl, undergraduate student in Imaging Science at the Rochester Institute of Technology, applied principal component analysis (PCA) to the three bands of the RGB color image of one Alexander leaf under ultraviolet illumination. PCA derives an equivalent orthogonal set of three bands from the original color image based on the image statistics and that are ordered by variance. Since the color image of the spectral fluorescence is dominated by the response in the blue color band that is closest to the excitation wavelength, the first principal component most closely resembles the blue channel of the original raw image. The second and third principal components are orthogonal to the first PC band and therefore are dominated by combinations (weighted sums) of the green and red color channels of the original raw file. The original text in different sections of the leaf is most visible in either the second or third principal component. These two bands are scaled to fill the available dynamic range, which enhances their relative contrast. A comparison to the original appearance for f. 120v is shown in Figure 4 [10].

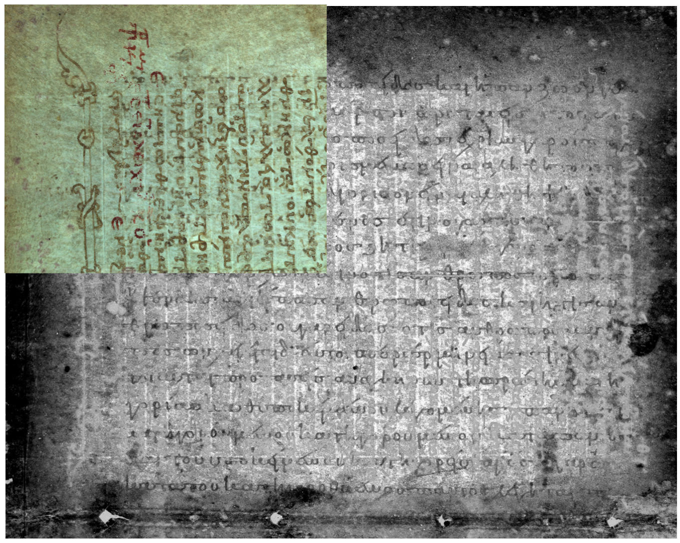


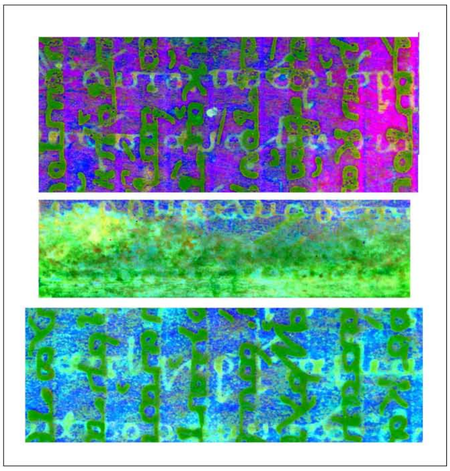
**Figure 4: Comparison of original appearance of f. 120v (upper left) to PCA band 3 from RGB image under UV illumination, showing original text visible in the latter. (Copyright retained by the owner of the Archimedes Palimpsest)**

### 6.3 HUE-ANGLE ROTATION OF PCA BANDS (2010)

A further improvement in the PCA technique was made in the summer of 2010 based on the fact that the text information is dominated by weighted combinations of the green and red channels of the fluorescence image . This suggested that pseudocolor renderings of the PC bands might be useful for scholarly readings. The three PC bands were mapped into the three color channels. The color contrast is enhanced because the dynamic ranges of the second and third PC bands are spread over the available monochrome bands. These images were presented to the scholars with instructions to dynamically rotate the hue angle of the pseudocolor image in software. This rotation of hue angle may improve the local visibility of the original text at different locations on the page. The scholars found these

renderings to be useful for completing the transcription of the Aristotle leaves. Much of the value of the dynamic rendering of pseudocolor images is not captured by static printed images, but Figure 5 attempts to illustrate by showing the pseudocolor image of three different sections of f.120v-121r (above, in, and below the gutter) at two different hue angles. Note the improved readability of the text in the gutter region at the second hue angle.

(a)



(b)

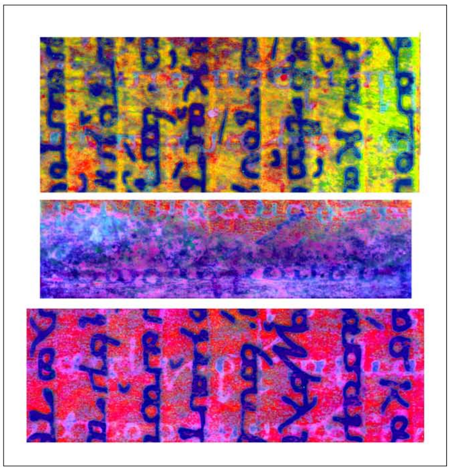


**Figure 5: Renderings of PCA pseudocolor image of f. 120v-121r that differ in hue angle. The top, middle, and bottom sections are respectively from the upper folio, gutter, and lower folio The renderings in (b) show additional text in the gutter. (Copyright retained by the owner of the Archimedes Palimpsest)**

## 7. CONCLUSIONS

A range of image collection protocols and processing tools were used to enhance the original texts in the 10th-century manuscript of the Archimedes Palimpsest. The processed images are providing a rich understanding of the sophistication of Archimedes' thinking. The readings from the other, previously unsuspected, original texts in the codex have also been judged to be of significance in philosophy and classical history.

Interested parties are invited and encouraged to apply their own methods to the raw and/or processed images, which are posted under a Creative Commons attribution license [2].


## 8. ACKNOWLEDGEMENTS

The authors are indebted to the owner of the Archimedes Palimpsest and to Dr. William Noel of the Walters Art Museum. Ken Boydston of Megavision, Inc.), the late John R. Stokes, and John T. Stokes of Stokes Imaging supplied imaging hardware and software essential to the task. We also thank the conservation staff of the Walters Art Museum, especially Abigail Quandt, for their invaluable contributions. We also thank Michael Toth for management expertise and Doug Emery for data support. Reviel Netz, Natalie Tchernetska, and Nigel Wilson, provided essential criticism. Michael Phelps of the Early Manuscripts Electronic Library and Dr. Fenella France of the U.S. Library of Congress provided guidance and perspective.

Matt Heimbueger, Allison Bright, Kevin Bloechl, Claire Mac Donald, and Teddy Hamlin performed image processing while students and/or interns at the Chester F. Carlson Center for Imaging Science.


## 9. REFERENCES


[1] Netz, R. and W. Noel (2007), ***The Archimedes Codex***, DaCapo Press, Cambridge.

[2] *http://www.archimedespalimpsest.net*

[3] Bergmann, U. and K.T. Knox (2009), *Pseudocolor enhanced x-ray fluorescence of the Archimedes Palimpsest*, **Proc. SPIE v.7247-02**.

[4] Griffin, C.W. (2007), *Digital Imaging: Looking Toward the Future of Manuscript Research*, **Currents in Biblical Research, v. 5**, 58-72.

[5] Schoengerdt, R.A. (2006), ***Remote Sensing, Models and Methods for Image Processing* (3rd *Edition)***, Academic Press, Burlingame, MA.

[6] Gippert, J. (2007), *The Application of Multispectral Imaging in the Study of Caucasian Palimpsests*, **Bul. Of Georgian Nat. Acad. Sciences, v. 175**, 168-179.

[7] Easton, R.L., Jr. and W.G. Noel (2004), *Multispectral imaging of the Archimedes Palimpsest*, **Gazette du Livre Médiéval, v. 45**, pp. 39-49.

[8] Easton, R.L., Jr. and W.G. Noel (2010), *Infinite Possibilities: Ten years of study of the Archimedes Palimpsest*, **Proc. Am. Philosophical Soc., v.154**, 50-76.

[9] Easton, R.L. Jr., K.T. Knox, W.A. Christens-Barry, K. Boydston, M.B. Toth, D. Emery., and W.G. Noel, (2010), *Standardized system for multispectral imaging of palimpsests*, **Proc. SPIE 7531-12**.

[10] Bloechl, K., H. Hamlin, and R. L. Easton, Jr. (2010), *Text recovery from the ultraviolet-fluorescence spectrum for a treatise in the Archimedes palimpsest*, **Proc. SPIE 7531-08**.